

A discusión

MODELLING SEGREGATION THROUGH CELLULAR AUTOMATA: A THEORETICAL ANSWER*

Juan Miguel Benito and Penélope Hernández**

WP-AD 2007-16

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.

Primera Edición Julio 2007

Depósito Legal: V-3154-2007

IVIE working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

* The authors thank partial financial support from the Spanish Ministry of Science and Technology under project SEJ2006-11510, from the Spanish Ministry of Education and Science under project SEJ2004-07554. We also wish to thank the Instituto Valenciano de Investigaciones Económicas (Ivie).

** J.M. Benito: Fundamentos del Análisis Económico I, Universidad del País Vasco-Euskal Eriko Unibertsitatea, Bilbao, Spain. E-mail address: jon.benito@ehu.es; P. Hernández: Departamento de Análisis Económico. Universidad de Valencia . Campus dels Tarongers. Edificio Departamental Oriental. Avda. dels Tarongers, s/n. 46022 Valencia. Spain. e-mail: penelope.hernandez@uv.es.

MODELLING SEGREGATION THROUGH CELLULAR AUTOMATA: A THEORETICAL ANSWER

Juan Miguel Benito and Penélope Hernández

ABSTRACT

This paper is a note in which we prove that Cellular Automata are suitable tools to model multi-agent interactive procedures. In particular, we apply the argument to validate results from simulation tools obtained for the classical model of segregation of Thomas Schelling (1971a).

Keywords: Cellular Automata, segregation, local information.

JEL Classification: C61, C73, D82.

1 Introduction

A deeper understanding of an economic system requires an understanding of how the individuals interact with each other, and how the results can be aggregated. A common characteristic of many of these multi-agent interactive procedures is the difficulty to find analytical outcomes. In the last decades, sophisticated techniques based on simulation tools, computational software, experimental approach, among others have permitted to represent and fit economic environments. *Agent-based Computational Economics* (ACE) concerns to the computational study of economies of such adaptive systems involving interacting agents (see Axelrod (1984)). More specifically, ACE seeks to understand how individuals behave and also the behavior of many individuals leading to large-scale outcomes. A growing proportion of ACE users produce computer simulations to construct and analyze the evolution over time of an economic world. There are many issues to which ACE's could be applied. For instance, a list of specific objectives are i) to find and explain global regularities such as social norms, ii) to discovery good designs to design economic policies, iii) to understand the potential dynamical behavior of an economic system under alternatively specified initial conditions, etc... Such understanding helps to clarify why certain non obvious consequences or global outcomes occur, even if the assumptions used to model the many interacting agents economic system are simple.

An important forerunner of the Economic literature which emphasizes the above issues was Thomas C. Schelling in his seminal papers (1969), (1971a). He studied the conditions under which individual residential location decisions interact to produce racially segregated neighborhoods. He assumed a population exhaustively divided into two groups; everyone is assumed to care about the color of the people where he lives among, and is able to observe the type distribution of the agents that occupy a piece of territory. Everyone has a particular location at any moment, and moreover he is capable of moving if he is dissatisfied with the color mixture of the location he is. Given this model, the first question which arises is how individual decisions of agents may affect the macro behavior of the system; and the second, how these decisions may organize themselves in space, in other words, what kind of spatial patterns may be observed on the global level.

The mentioned dynamics conveys a huge complexity degree and this is the reason why ACE becomes the most commonly used spatial model to obtain results on the Schelling model via Cellular Automata (see Albin (1998), Batten (2001), Epstein and Axtell (1996), Gaylord and D'Andria (1998), Laurie (2003), Páncs and Vriend (2007), Zhang (2004a), (2004b)). The ACE's simulation analysis considers Cellular Automata model in which the timing of updating is varied from synchronous to asynchronous. Cellular automata (CA) are dynamical systems defined on lattices in which time, states, and spatial relationships are discrete. The state of each cell on the

lattice updates at each time step according to a local rule which depends not only on the current state of the cell but also on the states of cells in its neighborhood around the cell. Up to now the CA approach has not been based on theoretical principles and then has not been enough to guarantee a rigorous conclusion even when the model insights many important results. Nevertheless, CA provide a simple, adaptable framework in which to analyze complex economic and social behavior through local interactions.

This paper presents a theoretical justification of a non-theoretical approach to ACE's simulation analysis with CA for the particular case of Schelling model. In other words, we provide a theoretical argument to validate conclusions from CA simulation analysis. Hence it could be considered as a methodology to apply and properly extend CA systems for many other economic environments.

We focus on the one-dimensional Schelling model of segregation for simplicity, but any other iterative problem with the same characteristics could also be tackled. For this purpose, we reformulate Schelling's dynamic as an algorithm that is defined as a list of simple and precise rules which stops in a finite number of steps. After the translation of the economic environment to the language of an algorithm, we use a famous and very accepted hypothesis called *Church-Turing thesis*, about the nature of mechanical calculation devices, such as electronic computers. The thesis claims that any possible calculation can be performed by an algorithm running on a computer, provided that both sufficient time and storage space are available. This implies that if we can express an economic dynamics by means of an algorithm then there will exist a program which could be run with our computer. Moreover, after a finite time (probably quite large) the computer will give an output.

The main result establishes the relationship between a Turing Machine and a Cellular Automata. Namely, for an arbitrary Turing Machine T with m symbols and n states, there exists a one-dimensional Cellular Automata A with a local rule with three neighbors and $m + 2n$ states which can simulate T . The consequences of the theorem are that Cellular Automata are suitable tools to study any phenomenon implemented by an algorithm, in particular, the Schelling model of segregation.

We address this question for the particular case of Schelling model following three steps: i) we express the Schelling model as an algorithm; ii) we implement it by a Turing Machine; iii) we prove the connection between Turing Machines and Cellular Automata.

The paper is organized as follows. Section 2 sets up the one-dimensional model of Schelling. The definitions of an algorithm and a Turing machine are refreshed in Section 3, the algorithm of the Schelling dynamics is presented as well. Our main result is offered in Section 4 where the connection between a Turing Machines and Cellular Automata is established.

2 The model of Schelling

There are two basic variants of Schelling's spatial proximity model (Schelling (1971a)), namely the one-dimensional and the two-dimensional. We focus on the first one without loss of generality¹. There are two types of agents, denoted by 0 and 1, uniformly distributed along a segment. An agent's position is defined relative to her neighbors only. At stage t , the society $S[t]$ is defined as a finite sequence of zeros and ones².

The Schelling's dynamics consists of three important ingredients: the first one is the information set of each agent which corresponds to her neighborhood of radio r ; the second one is a positive number $m \in \{1, \dots, 2r+1\}$ called the tolerance, which determines the maximum number of unlike neighbors that each agent is able to admit. In other words, tolerance could be understood as a threshold of dissatisfaction that each agent admits in her neighborhood; and finally, the individual utility which measures in a binary form the individual satisfaction level generated by her neighborhood. Formally, the information set of each agent i , her neighborhood of radio r at stage t denoted by $V^t(i, r)$, is equal to an element of $\{0, 1\}^{2r+1}$ centered on i and the utility of agent i at stage t is represented as follows:

$$U^t(i) = \begin{cases} 1 & \text{if } |\{j \in V^t(i, r) \text{ such that } j \neq i\}| \leq m \\ 0 & \text{if } |\{j \in V^t(i, r) \text{ such that } j \neq i\}| > m \end{cases}$$

The utility function says that each individual is concerned only with the number of like and unlike neighbors, this implies that agents care of the neighborhood's composition rather than of its configuration. More specifically, each agent wants at most m unlike neighbors; otherwise agents are indifferent.

The dynamic is an iterative process, where agents choosing myopic best-responses given agents' local information set. Specifically this is a sequential mechanism. At each stage, all unsatisfied agents are put in some arbitrary order. Schelling's movement arbitrarily let the discontented members to move in turn, counting from left to right. When it is her turn to move, each member will move to the nearest satisfactory location³, without regard if she had studied the prospective decisions of others whose turn comes later. Since all positions are relative only, she simply intrudes herself between two

¹Schelling (1969, 1971a, 1971b) also considers a two-dimensional variant as a regular lattice with bounds, such as a checkerboard. All the parameters and ingredients of this approach are the same that in the one-dimensional model. The main difference of this two-dimensional framework respect to the one-dimensional model is that agents can only move to empty positions.

²There exists a variant of Schelling model in which the possibility of an infinite continuous line or a ring is considered. Young (1998, 2001) also presents another variant of the Schelling's linear model in which agents are located in a ring.

³Nearest means the point reached by passing the smallest number of neighbors on the way.

agents (or either at the end of the line). Similarly, her own departure does not lead to an empty position. This process continues until no agent wants to move anymore.

The consequence of this dynamics is the emergence of a more segregated society than the initial one. Actually, two forces act in this process. The first one comes from agents' individual choice, whose depends on the composition of neighborhood rather than the configuration; while the second force is generated by the aggregation of the individual choices that determines the landscape of the society. The impact of the latter one is stronger, since although nobody actually prefers segregation to integration, the typical outcome is a highly segregated state of the society.

3 Implementation of the Schelling model by a Turing Machine

In this section we show that Schelling's dynamics can be implemented by a Turing Machine (henceforth TM), after rewritten such dynamics as an algorithm. Conceptually, a TM is an abstract mathematical model which formalizes the algorithm concept. In computation theory the *Church-Turing thesis*, named after Alonzo Church and Alan Turing, is an hypothesis about the nature of mechanical calculation devices, such as electronic computers. The thesis claims that any possible calculation can be performed by an algorithm running on a computer, provided that both sufficient time and storage space are available.

It is generally assumed that an algorithm must satisfy the following requirements:

1. The algorithm consists of a finite set of simple and precise instructions that are described by a finite number of symbols.
2. The algorithm will always produce the result in a finite number of steps.

Let us describe the features of the Schelling's dynamics by means of algorithm languages. Let $D[t]$ be the finite set of dissatisfied elements of $S[t]$ at stage t and $C[t]$ its cardinality. If $C[t] = 0$ then all members of the society are satisfied in their respective locations and the algorithm stops.

Suppose w.l.o.g. that $C[t] > 0$. Between two stages t to $t + 1$ there is a subroutine of $C[t]$ stages in which each agent in $D[t] = \{j(1), j(2), \dots, j(C[t])\}$ moves to his nearest place. Specifically, $j(1)$ is the first element from the left to the right at stage t in $S[t]$ who is not satisfied; $j(2)$ corresponds with the second dissatisfied agent, and so on. We can define inductively $S[t, k]$, such that $S[t, 1]$ is equal to $S[t]$, and $S[t, k + 1]$ is the society⁴ after the movement

⁴Notice that after some changes in $S[t]$, player $j(k)$ could decide to keep static.

of dissatisfied agent $j(k)$ in society $S[t, k]$. Therefore $S[t, C[t] + 1]$ corresponds to the final society when the subroutine is over, i.e. all agents in $D[t]$ have already moved.

The algorithm of Schelling's dynamic is as follows:

For each $t > 0$

If $C[t]=0$ then stop;

$S[t, 1]=S[t];$

For $k=1, \dots, C[t]+1$ $S[t, k+1]$

$S[t+1]=S[t, C[t]+1];$

If $S[t+1]$ in $S[1], \dots, S[t]$ then stop;

The above algorithm checks at any run: i) whether the society has no dissatisfied agents, i.e.: whether $D[t]$ is equal to the empty set for some $t \geq 1$; ii) whether $S[t+1]$ enters into a cycle, which will imply that the society never achieves a situation in which all individuals become satisfied. Under both situations the algorithm always stops. In the former (i) because the initial society is finite and in the later (ii) because of the algorithm's specification. Therefore, we capture all possible outcomes of Schelling's dynamics; either the society finishes with all members satisfied or the process never stops generating cycles.

4 From Turing Machines to Cellular Automata

In this section we provide a theoretical argument to validate conclusions from CA simulations. In Section 3 we have argued that any algorithm can be implemented by a TM. An one-dimensional Turing Machine is represented by a (finite) head and a (finite or infinite) tape of cells over finite alphabet. The machine reads an element of the tape, which corresponds with the input, and the head ascertains the output as a printed or erased symbol on the tape state to the left or to the right. A Cellular Automata are a finite one-dimensional linear array of cells of length. At every time, each cell has a value of a given alphabet. These values are updated iteratively according to a fixed rule which specifies exactly how the value of every site is computed from its own present value and the values of its immediate neighbors. Thus, the Cellular Automata have a head in each cell which acts synchronously.

Therefore, an one-dimensional Turing Machine is an abstract devise which can be adapted to simulate the *logic* of any computer that could be constructed and to study the properties of the list of simple rules call the algorithm. However, Turing Machines are not meant to be practical

computing technology, since they aren't actually constructed. On the other hand, Cellular Automata have been applied to both theoretical problems and experimental data analysis. In particular, Cellular Automata are suitable tools to simulate a certain class of phenomena. The question is which class of phenomena could be simulated by a Cellular Automata.

Our main result, Theorem 1, establishes the relationship between a Turing Machine and a Cellular Automata, and therefore allows us to assert that any phenomenon implemented by an algorithm can be studied by simulations with Cellular Automata. Formally, for an arbitrary one-dimensional Turing Machine T with m symbols and n states, there exists a one-dimensional Cellular Automata A with a local rule with three neighbors and $m + 2n$ states that can simulate it. The consequence of the theorem is that Cellular Automata are suitable tools to study any phenomenon implemented by an algorithm. In what follows we present the formal definitions of a TM and a CA and, the main theorem and its proof.

Definition 1 *A TM with an one-dimensional tape can be defined by the 6-tuple:*

$$TM = (Q, \Gamma, s, b, F, \delta)$$

where,

- Q is a finite set of states.
- Γ is the finite alphabet of tape symbols.
- $s \in Q$ is the start state.
- $b \in \Gamma$ is the blank symbol. The blank can appear infinite times.
- $F \subseteq Q$ is the set of final or accepting states.
- $\delta : Q \times \Gamma \rightarrow Q \times \Gamma \times \{L, R\}$ is the transition function, where L and R are the head moves toward left and right respectively.

Definition 2 *An one-dimensional CA can be defined by 5-tuple:*

$$CA = (\Gamma, S, r, \phi, c)$$

where,

- Γ is the finite alphabet of set symbol.
- S is the initial configuration of the one-dimensional linear array of length n , $S \in \Gamma^n$.
- $r \in \{1, \dots, n\}$ is a neighborhood ratio.

- ϕ is the CA rule, which assigns from a neighborhood configuration a new cell value, $\phi : \Gamma^{2r+1} \longrightarrow \Gamma$.
- c is the quiescent cell, which the CA rule leaves blocks of this kind of cells invariant, i.e., $\phi(c, \dots, c) = c$.

Table 1 gives a graphic illustration of this update process.

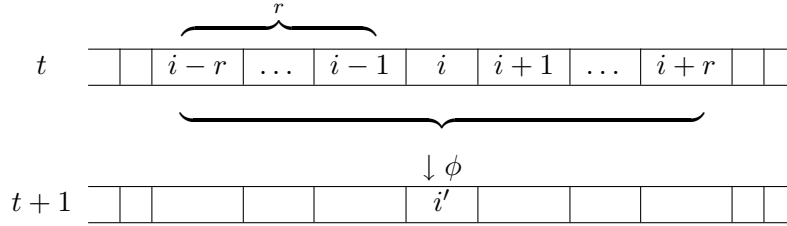


Table 1: The update process for a cell i in the CA lattice. The update rule ϕ is applied to the cell's local neighborhood configuration $(i-r, \dots, i, \dots, i+r)$ to determine the state of cell i at the next time step.

At each time step, all the cells in the lattice update their state *simultaneously* according to a CA rule. The CA's equation of motion is given by applying the rule to each point i separately, where the new value of cell in position i is given by:

$$i' = \phi(i-r, \dots, i, \dots, i+r).$$

Theorem 1 For an arbitrary one-dimensional Turing Machine T with m symbols and n states, there exists a one-dimensional Cellular Automata A with $r = 1$ (three neighbors) and with $m + 2n$ states that can simulate T .

Proof.

Let T be a Turing machine with a finite alphabet $\Gamma = \{x_1, x_2, \dots, x_m\}$ and $Q = \{q_1, q_2, \dots, q_n\}$ states. The movement of the head H is described by a set of quintuples of the form $q_i x_j x_k L q_l$ or $q_i x_j x_k R q_l$, where q_i is the intern state of T and q_l is the new intern state generated by the transition function; x_j corresponds to the symbol read from the tape and x_k is the one written in the tape; finally, L or R are the movements of the head.

We build an one-dimensional CA A that considers only the adjacent neighbor cells to the left (a_-) and to the right (a_+), i.e.: $r = 1$.

The linear array of cells represents the tape of a TM and each cell can be in any of the following $m + 2n$ states.

- The first m states $Q_M = \{x_1, x_2, \dots, x_m\}$ correspond to the m symbols of T .

- The states of T must be represented in A such that they specify the n states and also the movement of the head H ; in particular N states with movement toward the right $Q_{NR} = \{q_{1R}, \dots, q_{nR}\}$ and toward the left $Q_{NL} = \{q_{1L}, \dots, q_{nL}\}$, $Q_N = Q_{NL} \cup Q_{NR}$. Therefore we have $|Q_{NL} \cup Q_{NR}| = 2n$ states.
- The blank space in the tape of T is represented by quiescent cell.

Notice that the change of state occurs in a simultaneous way in A for each transition rule of T :

Transition in T	Transition in A				
	\sharp	a_-^t	a^t	a_+^t	a^{t+1}
$q_i x_j x_k R q_l$	1a	\dots	q_{iR}	x_j	x_k
	1b	q_{iR}	x_j	\dots	q_{lR}
$q_i x_j x_k L q_l$	2a	x_z	q_{iR}	x_j	q_{iL}
	2b	q_{iR}	x_j	\dots	x_k
	2c	\dots	x_z	q_{iL}	q_{lR}
	2d	x_z	q_{iL}	\dots	x_z

where (\dots) means that any state of the set Q_M , q_{iR}, q_{iL}, q_{lR} are in Q_N and x_i, x_k, x_z are in Q_M respectively. The movements in A toward the right side (1a and 1b of the preceding table) occur in one step:

$$\xrightarrow{t} x_1 x_2 q_R x_3 x_4 \xrightarrow{t+1} x_1 x_2 x'_3 q'_R x_4$$

and the movements toward the left side (2a, 2b, 2c, and, 2d) in two steps:

$$\xrightarrow{t} x_1 x_2 q_R x_3 x_4 \xrightarrow{t+1} x_1 x_2 q'_L x'_3 x_4 \xrightarrow{t+2} x_1 q'_R x_2 x'_3 x_4$$

The number of states of A is $|Q_M \cup Q_{NR} \cup Q_{NL}| = m + 2n$. ■

When r is the neighborhood ratio equal to 1, three neighbors, we can conclude that the model of Schelling can be simulated as a CA.

References

- [1] Albin, P.S. (1998). *Barriers and Bounds to Rationality: Essays on economic Complexity and Dynamics in Interactive Systems*. Princeton, New Jersey: Princeton University Press.
- [2] Axelrod, R. (1984). *The Evolution of Cooperation*. USA: Basic Books.
- [3] Batten, D.F. (2001). Complex landscapes of spatial interaction, *The Annals of Regional Science*, 35, 81-111.
- [4] Epstein, J. and Axtell, R. (1996). *Growing Artificial Societies*, Cambridge, MA: Brookings Institution Press and MIT Press.

- [5] Gaylord, R.J. and D'Andria, J. (1998). *Simulating Society*. New York: Springer-Verlag.
- [6] Laurie, J.A. and Jaggi, N.K. (2003). Role of 'Vision' in Neighbourhood Racial Segregation: A Variant of the Schelling Segregation Model. *Urban Studies*, 40(13), 2687-2704.
- [7] Pancs, R. and Vriend, N.J. (2007). Schelling's spatial proximity model of segregation revisited. *Journal of Public Economics*, 91, 1-24.
- [8] Schelling, T.C. (1969). Models of Segregation. *American Economic Review*, 59, 488-493.
- [9] Schelling, T.C. (1971a). Dynamic Models of Segregation. *Journal of Mathematical Sociology*, 1 (2), 143-186.
- [10] Schelling, T.C. (1971b). On the Ecology of Micromotives. *The Public Interest*, 25, 61-98.
- [11] Young, H.P. (1998). *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton, NJ: Princeton University Press.
- [12] Young, H.P. (2001). The Dynamics of Conformity. In S.N. Durlauf & H.P. Young (Eds.), *Social Dynamics*. Cambridge, Massachusetts: The MIT Press, 133-153.
- [13] Zhang, J. (2004). A dynamic model of residential segregation. *Journal of Mathematical Sociology*, 28, 147-170.
- [14] Zhang, J. (2004). Residential segregation in an all-integrationist world. *Journal of Economic Behavior & Organization*, 54, 533-550.